

Nazwa Wydziału	Wydział Matematyki i Informatyki
Nazwa jednostki prowadzącej moduł	Instytut Informatyki i Matematyki Komputerowej
Nazwa modułu kształcenia	<b>Wstęp do Przetwarzania Języka Naturalnego</b>
Kod modułu	WMI.II-WPJN-S
Język kształcenia	polski
Efekty kształcenia dla modułu kształcenia	<p>Wiedza</p> <ul style="list-style-type: none"> <li>• Student zna podstawowe pojęcia i metody przetwarzania języka naturalnego</li> <li>• posługuje się poprawną terminologią z tego obszaru</li> </ul> <p>Umiejętności</p> <ul style="list-style-type: none"> <li>• Student potrafi przeanalizować problem NLP i dobrać do niego odpowiednie podejście spośród znanych metod i modeli</li> <li>• Student potrafi zaimplementować odpowiednie rozwiązanie oraz zewaluować jego poprawność</li> </ul>
Typ modułu kształcenia (obowiązkowy/fakultatywny)	fakultatywny – ograniczonego wyboru (dla studentów danego instytutu)
Rok studiów	
Semestr	
Imię i nazwisko osoby/osób prowadzących moduł	mgr Wojciech Czarnecki
Imię i nazwisko osoby/osób egzaminującej/egzaminujących bądź udzielającej zaliczenia, w przypadku gdy nie jest to osoba prowadząca dany moduł	
Sposób realizacji	
Wymagania wstępne i dodatkowe	
Rodzaj i liczba godzin zajęć dydaktycznych wymagających bezpośredniego udziału nauczyciela akademickiego i studentów, gdy w danym module przewidziane są takie zajęcia	<p>Wykład, laboratorium</p> <p>Wykład: 30 Laboratorium: 30 <b>Łącznie: 60</b></p>
Liczba punktów ECTS	

przypisana modułowi	
Bilans punktów ECTS	
Stosowane metody dydaktyczne	<ul style="list-style-type: none"> <li>• Metody podające <ul style="list-style-type: none"> <li>○ wykład informacyjny</li> <li>○ objaśnienie lub wyjaśnienie</li> </ul> </li> <li>• metody praktyczne: <ul style="list-style-type: none"> <li>○ metoda projektów</li> <li>○ ćwiczenia przedmiotowe</li> </ul> </li> </ul>
Metody sprawdzania i kryteria oceny efektów kształcenia uzyskanych przez studentów	<p>Studenci są oceniani w sposób ciągły na podstawie uczestnictwa i aktywności na zajęciach oraz realizacji projektów.</p> <p>Metody kształtujące dla oceny ciągłej to:</p> <ul style="list-style-type: none"> <li>- bieżąca ocena i ewentualna korekta realizacji zadań wykonywanych w trakcie laboratoriów</li> <li>- ocena projektów na podstawie ich wydajności w odniesieniu do rozwiązywanego, praktycznego problemu</li> </ul> <p>Metody podsumowujące:</p> <ul style="list-style-type: none"> <li>- ostateczna, końcowa ocena wykonania zadań w tym zgodności z zaleceniami.</li> </ul>
Forma i warunki zaliczenia modułu, w tym zasady dopuszczenia do egzaminu, zaliczenia, a także forma i warunki zaliczenia poszczególnych zajęć wchodzących w zakres danego modułu	<p>Egzamin.</p> <p>Warunkiem dopuszczenia do egzaminu jest uzyskanie zaliczenia z laboratorium.</p> <p>Zaliczenie z laboratorium na ocenę.</p> <p>Zaliczenie na podstawie oddanych zadań (mini projektów) oraz aktywnego uczestnictwa.</p> <p>Obowiązuje następująca skala ocen (1-100 pkt.):</p> <ul style="list-style-type: none"> <li>• 0-50 pkt. – ocena ndst</li> <li>• 51-60 pkt. – ocena dst</li> <li>• 61-70 pkt. – ocena + dst</li> <li>• 71-80 pkt. – ocena db</li> <li>• 81-90 pkt. – ocena + db</li> <li>• 91-100 pkt. – ocena bdb</li> </ul>
Treści modułu kształcenia	<p>Celem zajęć jest zapoznanie studentów z podstawami analizy tekstu naturalnego. Zostaną przedstawione metody przetwarzania, analizy i rozumienia języka naturalnego (na podstawie języka angielskiego). Szczególny nacisk położony będzie na statystyczną analizę tekstu naturalnego, systemy uczące się, oraz stosowane współcześnie modele i algorytmy.</p> <p>W trakcie zajęć laboratoryjnych zostaną podane szczegóły techniczne poszczególnych rozwiązań oraz zostanie przedstawiony szereg narzędzi (w postaci bibliotek języka python) wspomagających tworzenie</p>

	<p>oprogramowania do analizy języka naturalnego. Studenci będą implementować poszczególne rozwiązania z nastawieniem na pracę własną (nacisk położony jest na realizację określonych zadań, nie zaś na użycie z góry narzuconej formy).</p> <p>Ramowy plan zajęć:</p> <ol style="list-style-type: none"> <li>1. Wprowadzenie do NLP</li> <li>2. Wyrażenia regularne</li> <li>3. Preprocessing tekstu, tokenizacja, lematyzacja, stemizacja</li> <li>4. Statystyczny model języka a klasyfikacja Bayesowska</li> <li>5. Ocena jakości statystycznych modeli języka</li> <li>6. Ukryte Modele Markova</li> <li>7. Modele generatywne i dyskryminatywne</li> <li>8. Tagowanie sekwencji</li> <li>9. Wektoryzacja dokumentów i miary ich podobieństwa</li> <li>10. Metody jądrowe</li> <li>11. Jądra tekstowe i semantyczne</li> <li>12. Nowoczesne metody analizy języka</li> </ol>
<p>Wykaz literatury podstawowej i uzupełniającej, obowiązującej do zaliczenia danego modułu</p>	<p>Literatura podstawowa:</p> <ol style="list-style-type: none"> <li>1. Daniel Jurafsky and James H. Martin (2008). <i>Speech and Language Processing</i>, 2nd edition. Pearson Prentice Hall. ISBN 978-0-13-187321-6.</li> <li>2. Steven Bird, Ewan Klein, and Edward Loper (2009). <i>Natural Language Processing with Python</i>. O'Reilly Media. ISBN 978-0-596-51649-9.</li> </ol>
<p>Wymiar, zasady i forma odbywania praktyk, w przypadku, gdy program kształcenia przewiduje praktyki</p>	